JOURNAL OF CLINICAL MICROBIOLOGY, Aug. 2005, p. 4112–4120 Vol. 43, No. 8
0095-1137/05/$08.00+0 doi:10.1128/JCM.43.8.4112–4120.2005
Copyright © 2005, American Society for Microbiology. All Rights Reserved.

Use of a Novel Panel of Nine Short Tandem Repeats for Exact and High-Resolution Fingerprinting of *Aspergillus fumigatus* Isolates

Hanneke A. de Valk,[1] Jacques F. G. M. Meis,[1] Ilse M. Curfs,[1] Konrad Muehlethaler,[2] Johan W. Mouton,[1] and Corné H. W. Klaassen[1]*

Department of Medical Microbiology and Infectious Diseases, Canisius Wilhelmina Hospital, Nijmegen, The Netherlands,[1] and Institute for Infectious Diseases, University of Bern, Bern, Switzerland[2]

Received 14 January 2005/Returned for modification 28 March 2005/Accepted 22 April 2005

Here we describe a new panel of short tandem repeats (STRs) for a novel exact typing assay that can be used to discriminate between *Aspergillus fumigatus* isolates. A total of nine STR markers were selected from available genomic *A. fumigatus* sequences and were divided into three multicolor multiplex PCRs. Each multiplex reaction amplified three di-, tri-, or tetranucleotide repeats, respectively. All nine STR markers were used to analyze 100 presumably unrelated *A. fumigatus* isolates. For each marker, between 11 and 37 alleles were found in this population. One isolate proved to be a mixture of at least two different isolates. With the remaining 99 isolates, 96 different fingerprinting profiles were obtained. The Simpson's diversity index for the individual markers ranged from 0.77 to 0.97. The diversity index for the multiplex combination of di-, tri-, and tetranucleotide repeats ranged from 0.9784 to 0.9968. The combination of all nine markers yielded a Simpson's diversity index of 0.9994, indicative of the high discriminatory power of these new loci. In theory, this panel of markers is able to discriminate between no less than 27×10^9 different genotypes. The multicolor multiplex approach allows large numbers of markers to be tested in a short period of time. The exact nature of the assay combines high reproducibility with the easy exchange of results and makes it a very suitable tool for large-scale epidemiological studies.

Aspergillus species are widely distributed fungi which release large amounts of conidia that are dispersed into the environment by air. Inhalation of *Aspergillus* conidia is a common event but rarely causes complications in immunocompetent hosts. In immunocompromised individuals, however, it can cause invasive aspergillosis (IA), often with fatal consequences. Because of the increase in the numbers of patients undergoing bone marrow or solid organ transplantation, the incidence of IA has increased dramatically in recent years. *Aspergillus fumigatus* is the most common species responsible for IA (9).

Molecular tools may provide a better understanding of the genetic and the epidemiological relationships between environmental and clinical isolates and thereby allow assessment of potential routes of transmission. Understanding of these routes may lead to improved strategies toward the prevention of *Aspergillus* infections. Several phenotypic and genotypic studies that can be used to distinguish between individual *A. fumigatus* isolates have been described. Phenotypic typing schemes include morphological and serological features as well as isoenzyme analysis. Genotypic methods include several techniques, such as random amplified polymorphic DNA (RAPD) analysis (1, 10, 11), restriction fragment length polymorphism (RFLP) analysis (13), amplified fragment length polymorphism (AFLP) analysis (18), and microsatellite length polymorphism (2). The last technique is sometimes also called microsatellite polymorphism analysis (5) or polymorphic microsatellites marker analysis (8). However, a major problem with pattern-based techniques, such as RAPD, RFLP, and AFLP analyses, is the poor interlaboratory reproducibility. The exchange of the results obtained by these techniques is therefore very difficult, if not impossible. Microsatellites, or short tandem repeats (STRs), are extensively being used for high-resolution fingerprinting of the human genome, but they have also been shown to provide a high level of discrimination between different isolates of several bacterial pathogens (7, 14, 16) and yeast (17). If they are properly performed, STR analyses yield highly reproducible, exact typing results. This would eliminate the need for repetitive analyses and would allow the easy exchange of data. We describe here a new panel of nine STRs for the exact and high-resolution fingerprinting of *A. fumigatus*.

(Part of this work was presented at the 44th Interscience Conference on Antimicrobial Agents and Chemotherapy [H. A. de Valk, J. F. G. M. Meis, I. M. Curfs, J. W. Mouton, and C. H. W. Klaassen, Abstr. 44th Intersci Conf. Antimicrob. Agents Chemother., abstr. K-1441, 2004].)

MATERIALS AND METHODS

Tandem repeat locus identification. Preliminary sequence data were obtained from The Wellcome Trust Sanger Institute (www.sanger.ac.uk) and were analyzed for the presence of STRs by using the Tandem Repeats Finder software (http://c3.biomath.mssm.edu/trf.html) (4). Since the *A. fumigatus* genome is estimated to contain about 35×10^6 bp, the available genomic sequences may represent nearly the entire genome. The minimum alignment score was set at 40, and the maximum period size (repeat unit) was set at 10 nucleotides. We selected three di-, tri-, and tetranucleotide repeats based on loci with the highest repeat numbers and counterselected on loci containing two or more repeat sequences within the boundaries of potential PCR primer regions. Interrupted repeats which may have a lower chance of displaying interstrain repeat number variation were not taken into consideration.

* Corresponding author. Mailing address: Department of Medical Microbiology and Infectious Diseases, Canisius Wilhelmina Hospital, Nijmegen, The Netherlands. Phone: 31-24-3657514. Fax: 31-24-3657516. E-mail: c.klaassen@cwz.nl.

TABLE 1. Overview of amplification primers for selected STR loci, details of repeat characteristics, and discriminatory indices

PCR and primer name	Forward primer sequence[a]	Reverse primer sequence[a]	Repeat unit	No. of repeats[b]			No. of alleles	*D* value[d]
				Minimum	Maximum	Ref[c]		
M2								
STR*Af* 2A	FAM-AAGGGTTATGGCCATTAGGG	GACCTCCAGGCAAAATGAGA	GA	10	28	26	14	0.88
STR*Af* 2B	HEX-TATTGGATCTGCTCCCAAGC	GAGATCATGCCCAAGGATGT	AG	9	29	18	16	0.88
STR*Af* 2C	TET-TCGGAGTAGTTGCAGGAAGG	AACGCGTCCTAGAATGTTGC	CA	8	33	18	17	0.88
M3								
STR*Af* 3A	FAM-GCTTCGTAGAGCGGAATCAC	GTACCGCTGCAAAGGACAGT	TCT	10	61	46	37	0.97
STR*Af* 3B	HEX-CAACTTGGTGTCAGCGAAGA	GAGGTACCACAACACAGCACA	AAG	8	73	20	14	0.84
STR*Af* 3C	TET-GGTTACATGGCTTGGAGCAT	GTACACAAAGGGTGGGATGG	TAG	5	51	23	32	0.86
M4								
STR*Af* 4A	FAM-TTGTTGGCCGCTTTTACTTC	GACCCAGCGCCTATAAATCA	TTCT	5	24	11	14	0.84
STR*Af* 4B	HEX-CGTAGTGACCTGAGCCTTCA	GGAAGGCTGTACCGTCAATCT	CTAT	5	26	10	11	0.80
STR*Af* 4C	TET-CATATTGGGAAACCCACTCG	ACCAACCCATCCAATTCGTAA	ATGT	5	45	8	15	0.77

[a] All primer sequences are given in the 5′ to 3′ direction.
[b] Number of repeats found in the reference population described in this study.
[c] Number of repeats in the reference strain AF293, upon which the genome sequences available from the *A. fumigatus* Sequencing Group at the Sanger Institute are based.
[d] The calculated *D* value according to Simpson's index of diversity.

Isolates. Ninety-nine clinical *A. fumigatus* isolates and one reference strain (CBS 487.65) were used in this study. The isolates were assumed to be unrelated because they were all collected from different patients either at different hospitals, from different wards, or at different points in time. The isolates were collected at six different hospital centers; five are located in The Netherlands, and one is located in Bern, Switzerland. Fifteen isolates were collected from patients with IA over a period of 4 years obtained from the University Hospital Nijmegen (Nijmegen, The Netherlands; location A). Another 29 isolates were obtained from the Canisius Wilhelmina Hospital (Nijmegen, The Netherlands); of these 29 isolates, 16 were from patients on the Intensive Care Unit (location B) and 13 were from patients the Pulmonology Ward (location C). Eleven isolates were collected over a period of 2 years from a cystic fibrosis center (Dekkerswald, Nijmegen, The Netherlands; location D). Another 14 isolates were from the Maashospital (Boxmeer, The Netherlands; location E). Eight isolates were from the University Hospital Maastricht (Maastricht, The Netherlands; location F), and 22 isolates were collected at the University Hospital Bern (Bern, Switzerland; location G). Isolates were stored at −80°C, according to standard procedures. Eleven *Aspergillus* species other than *A. fumigatus* were taken from the private collection at the Canisius Wilhelmina Hospital.

Identification of isolates. Fungal isolates were identified by their macroscopic and microscopic characteristics. Furthermore, all *A. fumigatus* isolates were tested for their ability to grow at 48°C.

DNA isolation. DNA was isolated by using the following culture and pretreatment conditions. Isolates were grown on Sabouraud agar plates at 35°C until sporulation. A prewetted cotton swab was saturated with conidia from a sporulating culture. Next, the spores were resuspended in a vial containing 350 μl lysis buffer and ceramic beads (Roche Diagnostics, Almere, The Netherlands) and subjected to mechanical lysis in a Magnalyzer instrument (Roche Diagnostics) for 30 s at 6,500 rpm. Following pretreatment, the DNA was further extracted and purified with a MagNA Pure LC instrument (Roche Diagnostics) in combination with a MagNA Pure DNA isolation kit III, according to the recommendations of the manufacturer. This DNA extraction protocol routinely yielded 5 to 20 μg DNA of excellent purity, as determined by UV absorbance measurements.

PCR amplification and genotyping. Three separate multiplex PCRs (M2, M3, and M4, respectively), each of which contained three different STRs, were developed. Multiplex PCR 2 (M2) amplified three dinucleotide loci. Multiplex PCR 3 (M3) amplified three trinucleotide loci, and multiplex PCR 4 (M4) contained three tetranucleotide loci. For all loci selected, PCR primers were designed with the use of the Primer3 software (http://frodo.wi.mit.edu/cgi-bin/primer3/primer3_www.cgi) (15). For each of the three multiplex PCRs, the forward primers were labeled with carboxyfluorescein (FAM), hexachlorofluorescein (HEX), or tetrachlorofluorescein (TET) at the 5′ end, respectively. An overview of the PCR primer sequences is shown in Table 1. Each PCR mixture contained 1 μM of all amplification primers, 0.2 mM deoxynucleoside triphosphates, 1 U of FastStart *Taq* DNA polymerase (Roche Diagnostics), and 1 ng of genomic DNA in 1× reaction buffer. The optimal $MgCl_2$ concentrations were 1.5 mM for M2 and 3.0 mM for M3 and M4. Thermocycling was performed in a T1 thermocycler (Biometra, Göttingen, Germany) by using the following thermal protocol: 10 min of denaturation at 95°C, followed by 30 cycles of 30 s of denaturation at 95°C, 30 s of annealing at 60°C, and 1 min of extension at 72°C. Before the reaction mixtures were cooled to room temperature, an additional incubation for 10 min at 72°C was performed. All temperature transitions were performed with maximal heating and cooling settings (5°C/s). The fragments obtained were combined with the ET400-R size standard (GE Healthcare, Roosendaal, The Netherlands) and analyzed on a MegaBACE 500 automated DNA platform (GE Healthcare), according to the instructions of the manufacturer. In order to avoid PCR amplicon contamination, pre- and post-PCR procedures (DNA isolation, preparation of master mixes, amplicon analysis) were performed in physically separated facilities.

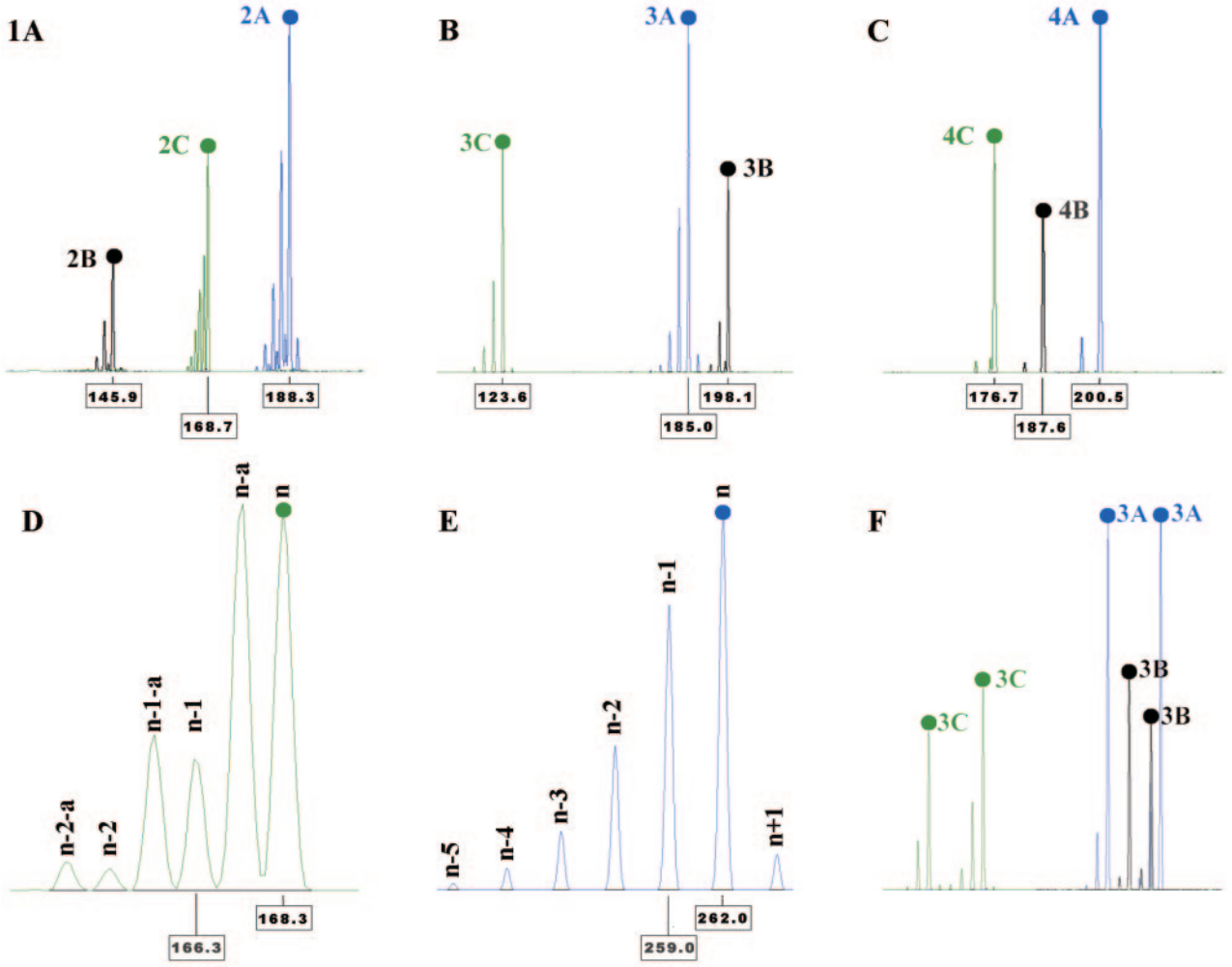
DNA sequence analysis. To determine the exact number of repeats in the obtained PCR products, a selected number of fragments (see Table 3) were inserted into the pGEM-T Easy Vector (Promega, Leiden, The Netherlands). After transformation of competent JM109 cells, individual colonies were selected and grown in liquid Luria-Bertani medium under selective pressure. Plasmid DNA was purified from the cultures that were obtained by using High Pure chemistry (Roche Diagnostics) and were sequenced with a universal M13 primer by using the MegaBACE DYEnamic ET Dye terminator kit, as suggested by the manufacturer (GE Healthcare). The reaction products were purified by using SPRI chemistry (GC-Biotech, Schiedam, The Netherlands), eluted in distilled water, and analyzed on a MegaBACE 500 capillary DNA analysis platform (GE Healthcare) by using standard electrophoretic conditions.

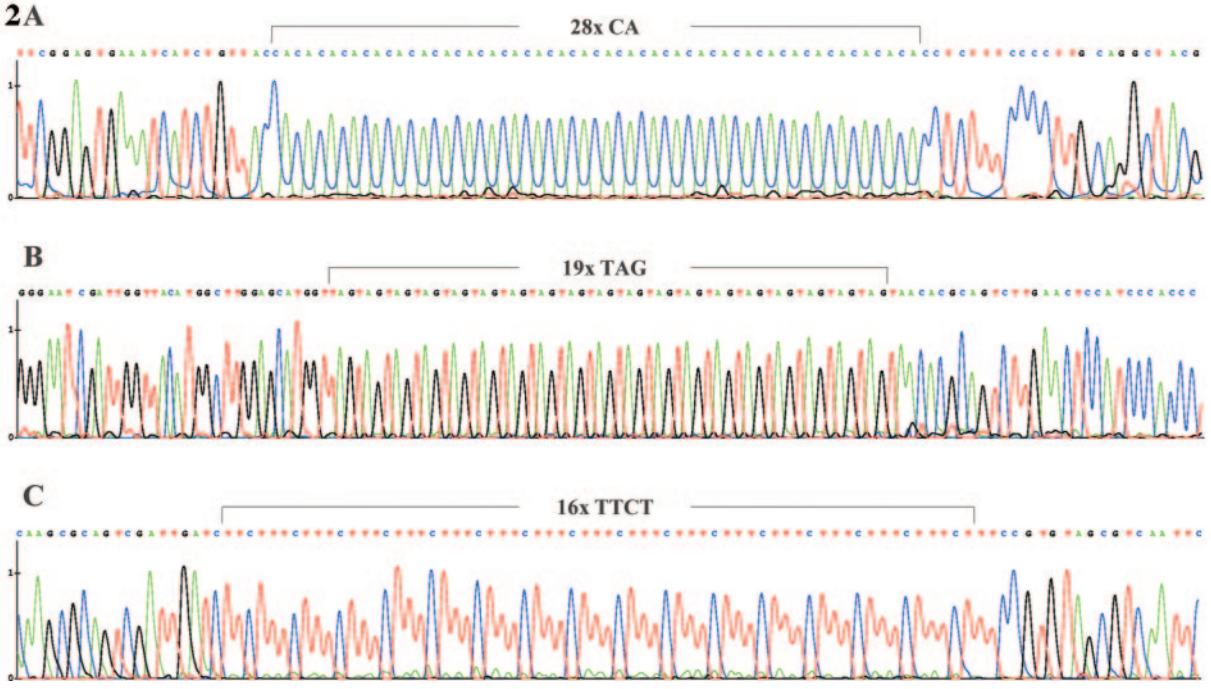
Data analysis. The repeat numbers of the nine markers of all isolates were analyzed by using BioNumerics, version 3.5, software (Applied Maths, Kortrijk, Belgium) and the unweighted pair group method with arithmetic averages with the multistate categorical similarity coefficient. All markers were given an equal weight. In the resulting dendrogram, the indicated percentages reflect the number of corresponding markers. Accordingly, two strains with six of the nine corresponding markers are 66.7% identical.

Discriminatory power. The discriminatory power of a typing method can be defined mathematically as the probability that two unrelated isolates chosen at random from a test population will be different by that typing method and can be calculated by using the Simpson index of diversity (*D*):

$$D = 1 - \frac{1}{N(N-1)} \sum_{j=1}^{s} n_j(n_j - 1)$$

where N is the total number of isolates in the test population, s is the total number of types described, and n_j is the number of isolates belonging to the jth type (6). A *D* value of 1.0 indicates that the typing method is able to discriminate between all isolates. A *D* value of 0.0 indicates that all isolates are identical.


1A
2A
2C
2B
145.9
168.7
188.3
B
3A
3C
3B
123.6
185.0
198.1
C
4A
4C
4B
176.7
187.6
200.5
D
n-a
n
n-1-a
n-1
n-2-a
n-2
166.3
168.3
E
n
n-1
n-2
n-3
n-4
n-5
n+1
259.0
262.0
F
3A
3A
3C
3C
3B
3B


2A
28x CA
B
19x TAG
C
16x TTCT

RESULTS

A selection of nine loci (three dinucleotide repeats, three trinucleotide repeats, and three tetranucleotide repeats) was made, and three multiplex PCRs (M2, M3, and M4) were developed and optimized. These three multiplex PCRs were used to analyze 100 *A. fumigatus* isolates from nonoutbreak situations. Typical amplification results are shown in Fig. 1.

The PCR amplification products of the tetranucleotide repeats consisted mainly of a single peak. However, the amplification products of the tri- and dinucleotide repeats usually contained more peaks, as they were one to three repeat units shorter. However, the longest peak was usually the most intense peak; this one was used for analysis (Fig. 1). Under some conditions (especially with high DNA concentrations), an extra peak 1 bp shorter than the main peak was observed; this is another established PCR artifact (Fig. 1).

In order to determine the exact repeat number of the repeats in all markers, a selected number of alleles throughout the entire range of alleles obtained were analyzed by DNA sequence analysis (Fig. 2). The repeat numbers of the remaining alleles (being at regularly spaced intervals) were determined by extrapolation, so that an exact repeat number could be specified for all nine loci.

The collection of 100 isolates was analyzed by using the complete panel of nine STR markers. One isolate displayed multiple peaks for all nine markers, which was probably the result of the fact that the isolate was a mixture of two different *A. fumigatus* isolates. With the remaining 99 isolates, a total of 96 different marker combinations were obtained. Three pairs of apparently unrelated isolates yielded the same combination. For all nine loci and combinations thereof, a *D* value was calculated (Tables 1 and 2). The highest discriminatory power value for a single locus was obtained with STR*Af* 3A, with a *D* value of 0.97. With this marker no less than 37 different alleles were found in this population. An overview of all alleles obtained is shown in Table 3. The combination of all nine markers yielded a *D* value of 0.9994.

If the variation in length between the different alleles were the sole effect of variations in repeat number, then the sizes of all alleles would be expected to occur at regularly spaced intervals (i.e., a 2-bp interval for the dinucleotide repeats and 3- and 4-bp intervals for the tri- and tetranucleotide repeats, respectively). Incidentally, a peak was found at a position halfway to the next allele (i.e., in STR*Af* 4C between repeat numbers 7 and 8). DNA sequence analysis of this allele demonstrated that the fragment contained a 2-bp deletion. Analogous to the nomenclature for such alleles in the human situation, this allele was denominated 7.2 (the same size as 7 full repeats plus 2 bp). Other confirmed aberrations were named accordingly.

TABLE 2. Diversity indices for all multiplex combinations

Combination of markers	No. of profiles[a]	*D* value[b]
M2	60	0.9879
M3	86	0.9968
M4	51	0.9784
M2 + M3	94	0.9988
M2 + M4	76	0.9928
M3 + M4	95	0.9992
M2 + M3 + M4	96	0.9994

[a] The total number of isolates is 99.
[b] The calculated *D* value is based on Simpson's index of diversity.

To test the stability of each locus, a random *A. fumigatus* isolate was subcultured a minimum of 30 times. DNA was isolated from every fifth subculture, and all DNA samples were tested by using all nine markers. The exact same number of repeats compared to the number in the original sample was obtained for all nine markers, indicating that the repeats are sufficiently stable to allow epidemiological surveys (data not shown).

In order to test the specificities of all markers for *A. fumigatus*, DNA from the following 11 other *Aspergillus* species were tested with all nine markers: *A. clavatus*, *A. chevalieri*, *A. flavus*, *A. glaucus*, *A. niger*, *A. nidulans*, *A. ochraceus*, *A. sodowii*, *A. terreus*, *A. ustus*, and *A. versicolor*. No amplification products were obtained with any of these species, indicating that the markers have a high specificity for *A. fumigatus*. Interestingly, one nonpigmented *A. fumigatus* variant from our collection yielded an amplification product with a single marker only, leaving the possibility of discrimination between cryptic *A. fumigatus* species.

DISCUSSION

At this time sequencing and assembly of the entire *A. fumigatus* genome approaches completion. This provided the opportunity to search through already available genomic sequences for potential new microsatellite loci in silico by using software available in the public domain. We applied a number of criteria to select or deselect candidate loci. Obviously, a repeat sequence can be regarded as a repeat sequence only if the repeat number is equal to or greater than 2. Since the chance that interstrain variation will be displayed increases with the repeat number, we started by selecting potentially suitable loci at the high end. Only loci with perfect repeat

FIG. 1. Examples of amplification results. The principle peak used for analysis is indicated with a colored dot. Blue traces represent the FAM label, black traces represent the HEX label, and green traces represent the TET label. Typical results obtained with the dinucleotide repeats (A), trinucleotide repeats (B), and tetranucleotide repeats (C) are shown. (D) Typical example of amplification artifacts obtained with a dinucleotide repeat. The principle peak is indicated with an "n." A second peak exactly 1 bp shorter is the result of incomplete A addition (hence, n − a). A stutter peak one repeat shorter is indicated with n − 1. Other artifact peaks are indicated accordingly. (E) Multiple stutter peaks with a trinucleotide repeat do not influence the ability to identify the principle peak in a sample. (F) A typical result obtained with a mixed sample.

FIG. 2. Determination of repeat numbers by DNA sequencing. Examples are shown for a dinucleotide repeat (A), trinucleotide repeat (B), and tetranucleotide repeat (C).

TABLE 3. Overview of the alleles obtained and their frequency distribution in this reference population

Allele	Fragment size (bp)	Repeat no.	Frequency (%)
STR*Af*2A	163.1	10[a]	9
	164.9	11	1
	168.8	13	3
	170.8	14[a]	8
	176.6	17	1
	178.6	18	21
	180.5	19	4
	182.4	20	5
	184.3	21	6
	188.3	23	20
	190.2	24	3
	192.2	25[a]	8
	194.0	26	9
	198.1	28	1
STR*Af*2B	124.6	9	1
	126.5	10[a]	1
	130.3	12[a]	15
	134.2	14	2
	138.1	16[a]	13
	140.2	17	1
	142.1	18	2
	144.0	19[a]	12
	146.0	20[a]	21
	147.8	21	7
	149.8	22	1
	151.6	23[a]	11
	153.7	24	2
	155.6	25	8
	161.4	28	1
	163.6	29[a]	1
STR*Af*2C	154.8	8	23
	156.5	9[a]	3
	158.7	10	12
	160.5	11	13
	162.8	12	4
	164.6	13[a]	2
	166.4	14[a]	3
	168.3	15	16
	170.6	16[a]	7
	172.1	17	2
	174.5	18	2
	176.4	19[a]	6
	184.3	23[a]	1
	189.9	26	1
	195.9	29[a]	2
	198.0	30	1
	203.9	33[a]	1
STR*Af*3A	138.1	10	5
	141.3	11[a]	1
	147.3	13	1
	150.5	14	1
	153.8	15	3
	156.5	16	2
	159.7	17	2
	163.3	18	2
	166.5	19[a]	1
	172.1	21	1
	175.3	22[a]	2
	178.6	23	2
	182.2	24	1
	184.9	25	6
	187.8	26[a]	6
	191.3	27	7
	194.4	28	3
	197.3	29	5
	200.7	30[a]	3
	203.5	31	5
	206.9	32	4
	209.9	33	2
	212.3	34	6
	215.9	35	4
	218.7	36	3
	222.0	37	4
	225.2	38	2
	228.4	39	2
	230.7	40[a]	2
	249.8	46[a]	1
	253.2	47	1
	255.5	48	2
	258.9	49	2
	262.0	50[a]	2
	284.1	57[a]	1
	290.2	59	1
	296.4	61	1
STR*Af*3B	158.0	8	1
	160.5	9	23
	163.5	10	21
	166.5	11	23
	169.3	12	7
	172.4	13	11
	177.9	15	1
	183.6	17	4
	186.6	18	1
	198.0	22[a]	2
	203.9	24	2
	226.9	32	1
	232.0	34[a]	1
	348.9	73	1
STR*Af*3C	77.5	5[a]	1
	80.1	6	3
	82.9	7	33
	86.2	8	1
	89.2	9[a]	1
	92.2	10	6
	98.3	12	1
	101.2	13	3
	104.6	14[a]	3
	107.5	15	1
	110.9	16	1
	114.0	17	2
	117.2	18	1
	120.5	19[a]	1
	123.4	20	11
	126.4	21	12
	129.5	22	1
	135.7	24[a]	1
	138.9	25	2
	144.9	27	1
	157.4	31	1
	160.4	32	1
	163.7	33	1
	166.9	34	1
	185.2	40[a]	1
	190.8	42	1
	200.5	45	1
	203.5	46	1
	206.9	47[a]	1
	212.5	49[a]	2
	215.8	50	1
	218.8	51	1

Continued on facing page

TABLE 3—*Continued*

Allele	Fragment size (bp)	Repeat no.	Frequency (%)	Allele	Fragment size (bp)	Repeat no.	Frequency (%)
STR*Af*4A	167.8	5[a]	2		196.0	12[a]	1
	176.2	7	9		236.9	22	1
	180.6	8	30		240.7	23	1
	185.0	9	10		249.5	25	1
	188.9	10[a]	21		252.9	26	9
	197.4	12	2				
	200.6	12.3[a]	5	STR*Af* 4C	164.0	5[a]	41
	201.4	13	9		168.0	6	4
	209.6	15[a]	3		172.5	7	4
	216.7	16.3	1		174.8	7.2[a]	1
	221.1	17.3	1		176.4	8	23
	226.3	19[a]	1		185.1	10	5
	233.8	21[a]	4		189.1	11	1
	246.3	24	1		193.3	12	2
					201.6	14[a]	2
STR*Af*4B	166.7	5[a]	9		213.7	17[a]	2
	175.3	7	1		225.8	20	9
	179.4	8	7		246.8	25[a]	1
	183.0	9	27		266.9	30[a]	2
	187.6	10[a]	32		295.8	37	1
	191.6	11	10		328.1	45	1

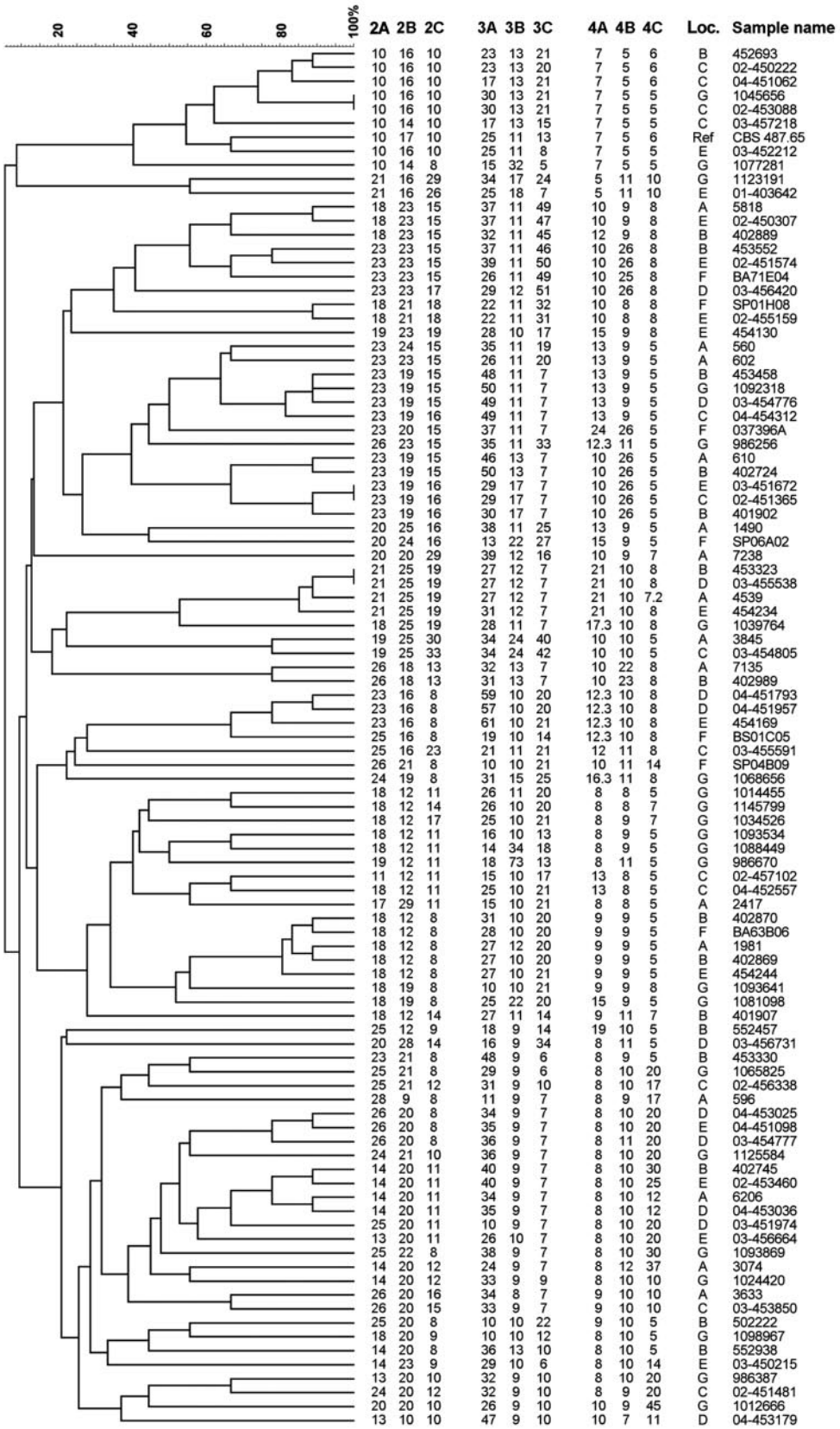
[a] The repeat was sequenced to determine the exact number of repeats; the repeat numbers of the remaining alleles were determined by extrapolation.

sequences were included in our analysis; imperfect repeats containing point mutations and/or insertion or deletions were excluded. Another criterion used for the exclusion of loci is the presence of a second repeat sequence. Certain loci with high repeat numbers were not considered due to the presence of an additional repeat sequence within the boundaries of potential PCR primer binding sites. Different combinations of two repeat sequences in a single PCR amplicon may lead to the formation of fragments of the same length. This makes it impossible to determine the exact repeat number for each of the two repeats and compromises the exact nature of the STR assay. The final exclusion criterion involved the absence of flanking sequences. Obviously, in these cases no PCR primer sequences can be designed. For this reason, a number of potentially interesting hexanucleotide repeat sequences had to be excluded. Once the entire genome sequence of *A. fumigatus* has been assembled, these loci also could prove to be interesting candidates for analysis. We finally selected three dinucleotide loci, three trinucleotide loci, and three tetranucleotide loci for further analysis.

Analysis of 100 presumed unrelated *A. fumigatus* isolates yielded 96 different profiles; 3 profiles were found twice, and in one sample double bands were obtained for all nine markers. As mentioned above, all isolates were assumed to be unrelated because they were all collected from different patients either at different hospitals, from different wards, or at different points in time. However, our results may indicate that the three profiles found twice were actually from related isolates. Unfortunately, studies comparing different fingerprinting techniques for *A. fumigatus* show that, up to now, no single technique can be considered the "gold standard" (3, 8). Additional fingerprinting of these isolates by AFLP analysis (18) also showed that these three pairs of isolates were indistinguishable from each other (results not shown). Since the genome of *Aspergillus* is haploid, double bands are typically a result of having started with an impure culture (i.e., a mixture of two different *A. fumigatus* isolates). In this case, the ratio between the heights and/or areas of these peaks will reflect the ratio between the two different isolates. The one isolate containing double bands for all nine markers was recultivated on Sabouraud agar plates. At the first occurrence of microcolonies, these were individually expanded on new agar plates and analyzed by using all nine markers. In these samples, two different profiles were obtained, each of which contained single peaks for all nine markers and each peak of which represented either of the two original peaks (not shown). The ability to easily identify the presence of multiple genotypes in a sample provides STR analysis with a unique advantage over other PCR fingerprinting methods like RAPD or AFLP analysis, by which such events usually remain unrecognized.

Unlike what might be expected from a fingerprinting technique that covers large parts of an organism's genome, this novel assay based on nine STR markers is probably not suitable for determination of the geographical origins of isolates. Indeed, there were no obvious subclusters of isolates from any of the hospitals (Fig. 3). This, however, must be substantiated by analyzing many more isolates of diverse geographical origins.

STR analysis is prone to several PCR artifacts that may influence the results. One of them is slippage of the *Taq* polymerase during amplification, a process that leads to the formation of so-called stutter peaks. There is a clear inverse relation between the size of the repeat unit and the formation of stutter peaks. With shorter repeat units, higher proportions of stutter peaks are generated (12). This is also clearly observed in our results, where stutter peaks are most prominently seen with the dinucleotide repeats, fewer stutter peaks are seen with trinucleotide repeats, and stutter repeat are almost absent with the tetranucleotide repeats. The occurrence of too many stutter peaks may complicate identification of the principle



2A	2B	2C	3A	3B	3C	4A	4B	4C	Loc.	Sample name
10	16	10	23	13	21	7	5	6	B	452693
10	16	10	23	13	20	7	5	6	C	02-450222
10	16	10	17	13	21	7	5	6	C	04-451062
10	16	10	30	13	21	7	5	5	G	1045656
10	16	10	30	13	21	7	5	5	C	02-453088
10	14	10	17	13	15	7	5	5	C	03-457218
10	17	10	25	11	13	7	5	6	Ref	CBS 487.65
10	16	10	25	11	8	7	5	5	E	03-452212
10	14	8	15	32	5	7	5	5	G	1077281
21	16	29	34	17	24	5	11	10	G	1123191
21	16	26	25	18	7	5	11	10	E	01-403642
18	23	15	37	11	49	10	9	8	A	5818
18	23	15	37	11	47	10	9	8	E	02-450307
18	23	15	32	11	45	12	9	8	B	402889
23	23	15	37	11	46	10	26	8	B	453552
23	23	15	39	11	50	10	26	8	E	02-451574
23	23	15	26	11	49	10	25	8	F	BA71E04
23	23	17	29	12	51	10	26	8	D	03-456420
18	21	18	22	11	32	10	8	8	F	SP01H08
18	21	18	22	11	31	10	8	8	E	02-455159
19	23	19	28	10	17	15	9	8	E	454130
23	24	15	35	11	19	13	9	5	A	560
23	23	15	26	11	20	13	9	5	A	602
23	19	15	48	11	7	13	9	5	B	453458
23	19	15	50	11	7	13	9	5	G	1092318
23	19	15	49	11	7	13	9	5	D	03-454776
23	19	16	49	11	7	13	9	5	C	04-454312
23	20	15	37	11	7	24	26	5	F	037396A
26	23	15	35	11	33	12.3	11	5	G	986256
23	19	15	46	13	7	10	26	5	A	610
23	19	15	50	13	7	10	26	5	B	402724
23	19	16	29	17	7	10	26	5	E	03-451672
23	19	16	29	17	7	10	26	5	C	02-451365
23	19	16	30	17	7	10	26	5	B	401902
20	25	16	38	11	25	13	9	5	A	1490
20	24	16	13	22	27	15	9	5	F	SP06A02
20	20	29	39	12	16	10	9	7	A	7238
21	25	19	27	12	7	21	10	8	B	453323
21	25	19	27	12	7	21	10	8	D	03-455538
21	25	19	27	12	7	21	10	7.2	A	4539
21	25	19	31	12	7	21	10	8	E	454234
18	25	19	28	11	7	17.3	10	8	G	1039764
19	25	30	34	24	40	10	10	5	A	3845
19	25	33	34	24	42	10	10	5	C	03-454805
26	18	13	32	13	7	10	22	8	A	7135
26	18	13	31	13	7	10	23	8	B	402989
23	16	8	59	10	20	12.3	10	8	D	04-451793
23	16	8	57	10	20	12.3	10	8	D	04-451957
23	16	8	61	10	21	12.3	10	8	E	454169
25	16	8	19	10	14	12.3	10	8	F	BS01C05
25	16	23	21	11	21	12	11	8	C	03-455591
26	21	8	10	10	21	10	11	14	F	SP04B09
24	19	8	31	15	25	16.3	11	8	G	1068656
18	12	11	26	11	20	8	8	5	G	1014455
18	12	14	26	10	20	8	8	7	G	1145799
18	12	17	25	10	21	8	9	7	G	1034526
18	12	11	16	10	13	8	9	5	G	1093534
18	12	11	14	34	18	8	9	5	G	1088449
19	12	11	18	73	13	8	11	5	G	986670
11	12	11	15	10	17	13	8	5	C	02-457102
18	12	11	25	10	21	13	8	5	C	04-452557
17	29	11	15	10	21	8	8	5	A	2417
18	12	8	31	10	20	9	9	5	B	402870
18	12	8	28	10	20	9	9	5	F	BA63B06
18	12	8	27	12	20	9	9	5	A	1981
18	12	8	27	10	20	9	9	5	B	402869
18	12	8	27	10	21	9	9	5	E	454244
18	19	8	10	10	21	9	9	8	G	1093641
18	19	8	25	22	20	15	9	5	G	1081098
18	12	14	27	11	14	9	11	7	B	401907
25	12	9	18	9	14	19	10	5	B	552457
20	28	14	16	9	34	8	11	5	D	03-456731
23	21	8	48	9	6	8	9	5	B	453330
25	21	8	29	9	6	8	10	20	G	1065825
25	21	12	31	9	10	8	10	17	C	02-456338
28	9	8	11	9	7	8	9	17	A	596
26	20	8	34	9	7	8	10	20	D	04-453025
26	20	8	35	9	7	8	10	20	E	04-451098
26	20	8	36	9	7	8	11	20	D	03-454777
24	21	10	36	9	7	8	10	20	G	1125584
14	20	11	40	9	7	8	10	30	B	402745
14	20	11	40	9	7	8	10	25	E	02-453460
14	20	11	34	9	7	8	10	12	A	6206
14	20	11	35	9	7	8	10	12	D	04-453036
25	20	11	10	9	7	8	10	20	D	03-451974
13	20	11	26	10	7	8	10	20	E	03-456664
25	22	8	38	9	7	8	10	30	G	1093869
14	20	12	24	9	7	8	12	37	A	3074
14	20	12	33	9	9	8	10	10	G	1024420
26	20	16	34	8	7	9	10	10	A	3633
26	20	15	33	9	7	9	10	10	C	03-453850
25	20	8	10	10	22	9	10	5	B	502222
18	20	9	10	10	12	8	10	5	G	1098967
14	20	8	36	13	10	8	10	5	B	552938
14	23	9	29	10	6	8	10	14	E	03-450215
13	20	10	32	9	10	8	10	20	G	986387
24	20	12	32	9	10	8	9	20	C	02-451481
20	20	10	26	9	10	10	9	45	G	1012666
13	10	10	47	9	10	10	7	11	D	04-453179

FIG. 3. Dendrogram based on profiles of nine STR*Af* markers from 99 presumably unrelated *A. fumigatus* isolates. The dendrogram was created by using BioNumerics, version 3.5, software (Applied Maths) and the unweighted pair group method with arithmetic averages method with the multistate categorical similarity coefficient. See the text for further details.

peak in a sample. For this reason, mononucleotide repeats, which often yield as many as 10 to 15 stutter peaks, were not taken into consideration. It has been suggested that addition of dimethyl sulfoxide to the amplification reactions or replacement of *Taq* DNA polymerase by other thermostable polymerases with higher processivities might reduce the formation of stutter peaks (12). Such adaptations were also evaluated, but none of them yielded improved results (data not shown). A basic understanding of potential PCR artifacts upon amplification and analysis of the STR sequences (as illustrated in Fig. 1D and E) is desired, but we conclude that the presence of stutter peaks will not influence the ability to identify the major peak in a sample when repeat units of three and above are used (i.e., at least trinucleotide repeats). This is a clear advantage over previously reported STR assays for *A. fumigatus*, where only dinucleotide repeats were analyzed (2). Another well-known PCR artifact is the addition of an extra A residue to the 3′ end of the PCR product by *Taq* DNA polymerase (known as template-independent polymerase activity or extendase activity). This activity is most pronounced when the ultimate residue at the 3′ end of the PCR product is a pyrimidine; the presence of a purine leads to incomplete A-residue addition and to the formation of additional unwanted peaks. To promote maximal addition of this A residue, all reverse amplification primers were designed to begin with a purine.

The number of different alleles found in our reference population with each of the nine markers is given in Table 1. This means that in theory no less than 27×10^9 different combinations can be discriminated with this panel of nine markers. This is an improvement of close to 10^6-fold over that obtained with the previously reported STR panel (2). If we assume that all intermittent alleles (not present in our reference population) also exist, this number increases to an overwhelming 3×10^{13} possible combinations. However, since there is no known sexual reproduction of *A. fumigatus*, all nine markers do not inherit independently from parent to daughter. It is therefore unlikely that all possible combinations indeed occur. As a natural consequence, certain (combinations of) alleles will occur more frequently than others, a phenomenon that was also observed in our reference population. Each combination of these nine markers should therefore be considered a single haplotype. The total number of different haplotypes can be determined only experimentally, but since we already found 96 different haplotypes in a collection of 99 presumably unrelated isolates, this indicates that this number must be quite large.

This novel combination of nine STR markers for fingerprinting of *A. fumigatus* with the potential to resolve 3×10^{13} combinations of alleles may appear to be somewhat overdone. Since each of the three multiplex PCRs already has a very high discriminatory power, it may not always be necessary to analyze all nine markers in order to discriminate between certain isolates. One could consider starting with just the M3 combination (since this one has the highest discriminatory power, which is even higher than that of M2 and M4 combined) and to include M4 or M2 only when it is necessary. The additional value of including the third combination appears to be rather limited, but it could always serve as an additional backup.

In order to make full use of the potential of STR analysis, high-resolution analysis of the fragments obtained is a prerequisite. Standard agarose gel electrophoresis equipment usually does not yield sufficient resolution to discriminate between fragments that differ by as little as 2 bp, but at best, it could just suffice for the analysis of tri- and tetranucleotide repeats. Use of agarose gels also means that each marker must be analyzed individually since the different loci in our multiplex analysis overlap each other. We chose to make multiplex combinations using a multicolor approach, since this allows testing for multiple loci in one run, which greatly improves throughput capabilities and allows rapid analysis of large numbers of markers. Since high-resolution equipment like capillary-based or acrylamide-based electrophoresis platforms are increasingly finding their way into the laboratories, this seems to be the most logical way to go. Use of high-resolution equipment also allows even 1-bp insertions or deletions to be identified reliably. If this is not an option and one would still prefer to use agarose-based systems, it would make more sense to restrict STR analysis to repeat units of 10 and beyond. Such repeats have not been analyzed in this work.

In this study we describe a new, highly discriminatory PCR fingerprinting assay for *A. fumigatus* with a novel panel of nine STRs. The multicolor multiplex approach allows large numbers of markers to be tested in a short period of time. The exact nature of the assay combines high reproducibility with the easy exchange of results.

ACKNOWLEDGMENTS

Preliminary sequence data were obtained from The Wellcome Trust Sanger Institute (www.sanger.ac.uk).

The *Aspergillus fumigatus* Sequencing Consortium was funded by the National Institute of Allergy and Infectious Diseases, The Wellcome Trust, and the Fondo de Investicagiones Sanitarias.

REFERENCES

1. **Aufauvre-Brown, A., J. Cohen, and D. W. Holden.** 1992. Use of randomly amplified polymorphic DNA markers to distinguish isolates of *Aspergillus fumigatus*. J. Clin. Microbiol. **30:**2991–2993.
2. **Bart-Delabesse, E., J.-F. Humbert, E. Delabesse, and S. Bretagne.** 1998. Microsatellite markers for typing *Aspergillus fumigatus* isolates. J. Clin. Microbiol. **36:**2413–2418.
3. **Bart-Delabesse, E., J. Sarfati, J.-P. Debeaupuis, W. van Leeuwen, A. van Belkum, S. Bretagne, and J.-P. Latge.** 2001. Comparison of restriction fragment length polymorphism, microsatellite length polymorphism, and random amplification of polymorphic DNA analyses for fingerprinting *Aspergillus fumigatus* isolates. J. Clin. Microbiol. **39:**2683–2686.
4. **Benson, G.** 1999. Tandem repeats finder: a program to analyze DNA sequences. Nucleic Acids Res. **27:**573–580.
5. **Bertout, S., F. Renaud, R. Barton, F. Symoens, J. Burnod, M.-A. Piens, B. Lebeau, M.-A. Viviani, F. Chapuis, J.-M. Bastide, R. Grillot, M. Mallie, and The European Research Group on Biotype and Genotype of Aspergillus.** 2001. Genetic polymorphism of *Aspergillus fumigatus* in clinical samples from patients with invasive aspergillosis: investigation using multiple typing methods. J. Clin. Microbiol. **39:**1731–1737.
6. **Hunter, P. R., and M. A. Gaston.** 1988. Numerical index of the discriminatory ability of typing systems: an application of Simpson's index of diversity. J. Clin. Microbiol. **26:**2465–2466.
7. **Klevytska, A. M., L. B. Price, J. M. Schupp, P. L. Worsham, J. Wong, and P. Keim.** 2001. Identification and characterization of variable-number tandem repeats in the *Yersinia pestis* genome. J. Clin. Microbiol. **39:**3179–3185.
8. **Lasker, B. A.** 2002. Evaluation of performance of four genotypic methods for studying the genetic epidemiology of *Aspergillus fumigatus* isolates. J. Clin. Microbiol. **40:**2886–2892.
9. **Latge, J.-P.** 1999. *Aspergillus fumigatus* and aspergillosis. Clin. Microbiol. Rev. **12:**310–350.
10. **Leenders, A., A. van Belkum, S. Janssen, S. de Marie, J. Kluytmans, J. Wielenga, B. Lowenberg, and H. Verbrugh.** 1996. Molecular epidemiology of apparent outbreak of invasive aspergillosis in a hematology ward. J. Clin. Microbiol. **34:**345–351.
11. **Loudon, K. W., J. P. Burnie, A. P. Coke, and R. C. Matthews.** 1993. Application of polymerase chain reaction to fingerprinting *Aspergillus fumigatus* by random amplification of polymorphic DNA. J. Clin. Microbiol. **31:**1117–1121.

12. **Murray, V., C. Monchawin, and P. R. England.** 1993. The determination of the sequences present in the shadow bands of a dinucleotide repeat PCR. Nucleic Acids Res. **21:**2395–2398.
13. **Neuveglise, C., J. Sarfati, J. Latge, and S. Paris.** 1996. Afut1, a retrotransposon-like element from *Aspergillus fumigatus*. Nucleic Acids Res. **24:**1428–1434.
14. **Onteniente, L., S. Brisse, P. T. Tassios, and G. Vergnaud.** 2003. Evaluation of the polymorphisms associated with tandem repeats for *Pseudomonas aeruginosa* strain typing. J. Clin. Microbiol. **41:**4991–4997.
15. **Rozen, S., and H. Skaletsky.** 2000. Primer3 on the WWW for general users and for biologist programmers. Methods Mol. Biol. **132:**365–386.
16. **van Belkum, A., W. Melchers, C. Ijsseldijk, L. Nohlmans, H. Verbrugh, and J. Meis.** 1997. Outbreak of amoxicillin-resistant *Haemophilus influenzae* type b: variable number of tandem repeats as novel molecular markers. J. Clin. Microbiol. **35:**1517–1520.
17. **Verduyn-Lunel, F., L. Licciardello, S. Stefani, H. A. Verbrugh, W. J. G. Melchers, J. F. G. M. Meis, S. Scherer, and A. van Belkum.** 1998. Lack of consistent short sequence repeat polymorphisms in genetically homologous colonizing and invasive *Candida albicans* strains. J. Bacteriol. **180:**3771–3778.
18. **Warris, A., C. H. W. Klaassen, J. F. G. M. Meis, M. T. de Ruiter, H. A. de Valk, T. G. Abrahamsen, P. Gaustad, and P. E. Verweij.** 2003. Molecular epidemiology of *Aspergillus fumigatus* isolates recovered from water, air, and patients shows two clusters of genetically distinct strains. J. Clin. Microbiol. **41:**4101–4106.